Queueing Syst (2011) 67: 63–90
DOI 10.1007/s11134-010-9201-2

Dynamic control of a single-server system with abandonments

Douglas G. Down · Ger Koole · Mark E. Lewis

Received: 2 August 2010 / Revised: 18 October 2010 / Published online: 16 November 2010

Abstract In this paper, we discuss the dynamic server control in a two-class service system with abandonments. Two models are considered. In the first case, rewards are received upon service completion, and there are no abandonment costs (other than the lost opportunity to gain rewards). In the second, holding costs per customer per unit time are accrued, and each abandonment involves a fixed cost. Both cases are considered under the discounted or average reward/cost criterion. These are extensions of the classic scheduling question (without abandonments) where it is well known that simple priority rules hold.

The contributions in this paper are twofold. First, we show that the classic c–μ rule does not hold in general. An added condition on the ordering of the abandonment rates is sufficient to recover the priority rule. Counterexamples show that this condition is not necessary, but when it is violated, significant loss can occur. In the reward case, we show that the decision involves an intuitive tradeoff between getting more rewards and avoiding idling. Secondly, we note that traditional solution techniques are not directly applicable. Since customers may leave in between services, an interchange argument cannot be applied. Since the abandonment rates are unbounded we cannot apply uniformization—and thus cannot use the usual discrete-time Markov decision process techniques. After formulating the problem as a *continuous-time*

D.G. Down
Department of Computing and Software, McMaster University, 1280 Main Street West, Hamilton, ON L8S 4L7, Canada
e-mail: downd@mcmaster.ca

G. Koole
Faculty of Sciences, Department of Mathematics, VU University, Amsterdam, The Netherlands
e-mail: koole@few.vu.nl

M.E. Lewis (✉)
School of Operations Research and Information Engineering, Cornell University, 226 Rhodes Hall, Ithaca, NY 14853, USA
e-mail: mark.lewis@cornell.edu

Markov decision process (CTMDP), we use sample path arguments in the reward case and a savvy use of truncation in the holding cost case to yield the results. As far as we know, this is the first time that either have been used in conjunction with the CTMDP to show structure in a queueing control problem. The insights made in each model are supported by a detailed numerical study.

Keywords Priority rules · Dynamic programming · Control of queues

Mathematics Subject Classification (2000) 90B36 · 60K25 · 90C40

1 Introduction

In many service systems, a server (or servers) is faced with the task of serving impatient customers. While one may attempt to implement methods to decrease levels of impatience, at the end of the day, a fundamental decision that must be made at any point in time is: given a particular cost/reward structure and any information about the impatience of customers, where should the server direct its effort? In this paper we provide models that are seemingly simple extensions of classic scheduling problems to include customer impatience. Our results suggest that the server may need to weigh the relative costs/benefits of avoiding idleness (by letting too many customers abandon) against short-term revenue maximization/cost minimization concerns. The models that we consider consist of independent, Poisson arrival streams for each class of customer. There is a single server to serve both classes. Service times are exponentially distributed with rates that are independent of the customer's class. To this basic setup we add that all customers may abandon after an exponentially distributed period of time, with the abandonment rates allowed to be class dependent. Our goal is to provide an optimal server assignment policy, which we do under two settings:

1. For each customer successfully completed, a class-dependent reward is received.
2. Each queue has (linear) holding costs, and there is a class-dependent penalty for each customer that abandons.

In each case we consider the problem of maximizing expected discounted or average rewards or minimizing expected discounted or average costs over an infinite horizon.

It is well known that for the second case above, if there are no abandonments, then the c–μ rule is optimal (see [7]). In this paper, we show that this is not always true when abandonments are considered. In fact, there is a tension between losing future workload through abandonments (and thus creating excessive idling) and myopically reducing costs (through the c–μ rule). For appropriate combinations of parameters, there is no tension between these two factors, in which case an appropriately modified version of a c–μ rule is optimal. We identify such combinations. Note that such a tension cannot be captured in other approaches to server control. One can think of the problems of server assignment as generally being handled by examining three different regimes.

1. Overloaded regime. Here, a fluid model approach is applicable. For our model, Atar et al. [6] show that a form of c–μ rule is indeed optimal. In a system with

many customer classes and a single server, a simple rule that prioritizes the class with the largest value of the product of the holding cost and service rate divided by the abandonment rate is shown to be asymptotically optimal in minimizing the long-run average holding cost. In this case, there is always work for the server to do, so there is no need to limit server idleness.
2. Critically loaded regime. A diffusion model approach is applicable. References for such an approach are Ghamami and Ward [10], Harrison and Zeevi [13], and Tezcan and Dai [21]. All of these formulate the solution to a diffusion control problem which yields priority policies under conditions similar to those developed in our approach.
3. Underloaded regime. This is the regime of our analysis. It is not clear that *any* asymptotic approach is appropriate. As mentioned above, analyzing this model brings in the issue of server idleness and the resulting tension with cost reduction. Our work is the first that we are aware of on systems with abandonments in this regime.

A combination of the insights developed by these three different approaches should provide a clear(er) view on how to control a server faced with abandoning customers.

In addition to the above references, there are a few other related works. To put into context the issue of abandonments in call centre models, the reader should consult the comprehensive surveys of Aksin et al. [1] and Gans et al. [8]. Related work includes that of Argon et al. [2], who show that for a clearing system with abandonments, the policy that minimizes the number of abandonments is that which serves jobs with the shortest lifetime and shortest service time (assuming that they can be ordered this way). The performance of strict priority policies is studied in Iravani and Balcıoğlu [14], but no optimality results are obtained. In [22, 23] Ward and Glynn study single-class systems with abandonments. They show that under appropriate distributional assumptions, G/G/1 queues with balking and/or reneging can be approximated (i.e., there is appropriate convergence in heavy traffic) with a regulated Ornstein–Uhlenbeck process. While in our work the only possibilities after arrival are that a customer is either served or abandons, there is a line of work that attempts to compensate for potential abandonments in other manners. In [15] Koçağa and Ward study an admission control problem for a multiserver queue with a single class of customers who may abandon. In Armony et al. [5] customers are provided with delay estimates to influence their behavior, while in Armony and Maglaras [3, 4] a call-back option is proposed to allow potential abandonments to be contacted at a future point in time (when presumably servers are less busy). Note that their approaches and ours can be seen to be complementary.

The methodology that we use is that of Markov Decision Processes. We see our work as having two significant contributions in this area.

1. Due to the abandoning customers, *uniformization* (cf. [16]) is not possible (transition rates are unbounded). Thus we do our analysis in continuous time to allow us to deal with the unbounded rates (cf. [12]). In addition to showing how one can handle unbounded rates, we see novelty in using a continuous-time framework to show structural results.
2. In the course of our analysis, we truncate a multidimensional state space and let the truncation level go to infinity. Not only is this limiting approach of interest, we

show that if truncation is done in a smart manner, analysis is greatly simplified (or goes from intractable to tractable). For related work on this, see [9].

In addition to our analytic results, we supplement our work with several numerical studies that show that the price of not taking into account abandonment rates can be significant. These studies also suggest relative ranges of parameters (in particular, abandonment rates and either rewards or costs, according to the model), for which looking beyond a c–μ rule can lead to significant improvements.

The rest of the paper is organized as follows: a complete description of the queueing dynamics, optimality criterion and a proof that we can restrict attention to nonidling policies are shown in Sect. 2. The optimal control in both the reward and holding cost models is covered in Sect. 3. A detailed numerical study is provided in Sect. 4, while conclusions and some suggestions for future work are contained in Sect. 5.

2 Model and preliminaries

In this section we define the queueing dynamics, then discuss two criteria that we use for design; the first is one in which a fixed (type-dependent) reward is received for each customer successfully completed, and we term this the *reward model*. The second considers a combination of holding costs and penalties for each customer that abandons and is called the *holding cost model*. In each case, we show that it is sufficient to restrict attention to nonidling policies. Finally, we give the optimality equations for both criteria and show that a solution exists in each case.

2.1 Queueing dynamics and optimality criteria

Suppose that two stations are served by a single server. Customer arrivals to stations 1 and 2 occur according to independent Poisson processes with rates λ_1 and λ_2, respectively. We will also refer to arrivals to station i as class i customers. Customer service requirements are probabilistically the same in the sense that they are exponential with rate 1. Customers at station 1 (2) have limited patience and are only willing to wait an exponentially distributed amount of time with rate $\beta_1 > 0$ ($\beta_2 > 0$). That is to say that the abandonment rate in station 1 is $i\beta_1$ when there are i customers there. Service is preemptive, and customers in service may abandon. A priori (since the transition rates are unbounded) we are not assured that each Markov policy, say π, yields a regular Markov process. For more information along these lines, please see [11] or the comments on p. 187 in [12]. Regularity is guaranteed by showing that for the current models, Assumption A of the Appendix holds.

Suppose that the state space is $\mathbb{X} = \{(i, j) : i, j \in \mathbb{Z}^+\}$, where i (j) represents the current number of customers at station 1 (2). Let $N(t)$ be a counting process that counts the number of decision epochs by time t, and σ_n represent the time of the nth epoch. We seek a policy that describes where to place the server based on the current state and potentially the history of states and actions taken; a nonanticipating policy.

The finite horizon, discounted expected reward or cost (depending on the model) for a nonanticipating policy π is

$$v_{\alpha,t}^{\pi}(i,j) = \mathbb{E}_{(i,j)}^{\pi} \sum_{n=0}^{N(t)} e^{-\alpha\sigma_n} k(X_n, a_n) + \int_0^t \bigl[e^{-\alpha s}\, \mathbb{E}_{(i,j)}^{\pi}\bigl[h_1 Q_1(s) + h_2 Q_2(s)\bigr]\bigr]\, ds,$$

where $Q_m(s)$ is the number of customers at station m, $m = 1, 2$, and X_n and a_n represent the state of the system and the type of event seen at the time of the nth decision, respectively. The function $k(\cdot,\cdot)$ denotes the fixed reward or cost depending on which model is under consideration. That is to say that in the rewards model $h_1 = h_2 = 0$ and if σ_n represents a service completion at station ℓ, then $k(X_n, a_n) = R_\ell$. In the holding cost model if σ_n represents an abandonment from station ℓ, then $k(X_n, a_n) = P_\ell$ (it is zero otherwise). For $\alpha > 0$, the infinite-horizon discounted expected cost under policy π is $v_\alpha^\pi(i,j) := \lim_{t\to\infty} v_{\alpha,t}^\pi(i,j)$. The long-run average reward (cost) rate is $\rho^\pi(i,j) := \liminf_{t\to\infty} \frac{v_{0,t}^\pi(i,j)}{t}$ $(\limsup_{t\to\infty} \frac{v_{0,t}^\pi(i,j)}{t})$. Under either optimality criterion in the rewards model, we seek a policy π^* such that $w^{\pi^*}(i,j) = \sup_{\pi\in\Pi} w^\pi(i,j)$, where Π is the set of all nonanticipating policies, and $w = v_\alpha$ or ρ. There is the obvious analogue in the holding cost model.

We end this section with the following preliminary result. It states the intuitive observation that it is better to have more customers in the system in the reward model and less in the system in the holding cost model. The proof is simple and is omitted for brevity.

Proposition 2.1 *Let $y = v_\alpha$ or $v_{\alpha,t}$ (with $\alpha \ge 0$ or $\alpha > 0$) depending on the optimality criterion. For either the reward or holding cost model, the following inequalities hold:*

1. $y(i, j+1) \ge y(i,j)$,
2. $y(i+1, j) \ge y(i,j)$,

where in the finite-horizon case the result holds for all $t \ge 0$.

2.2 Optimality of nonidling policies

In this section, we show that it suffices to consider only nonidling policies.

Proposition 2.2 *In either the reward or holding cost model, and under the finite horizon discounted cost criterion for any fixed and finite $t \ge 0$ and $\alpha \ge 0$, there exists an optimal policy that does not idle except when the system is empty.*

Proof We show the result for the reward model by showing how one can construct a nonidling policy that dominates one that idles. This is done via a sample path argument. The holding cost model is analogous (and is in fact simpler). Suppose that we start two processes on the same probability space, each starting in state (i,j) with $i \ge 1$. Suppose that Process 1 uses a policy ϕ that initially idles the server. Process 2 uses a policy $\tilde{\phi}$ that has the server working at station 1. If no events occur before the end of the horizon, there is no difference in the rewards. Similarly, if Process 1 begins to work again before Process 2 has a service completion, assume that both processes use the same policy thereafter and there is no difference in the expected reward stream.

Suppose now that Process 2 completes a service before the time horizon ends (at time, say, x) and before Process 1 begins working again. The difference in the total rewards is $v^{\phi}_{\alpha,t-x}(i', j') - R_1 - v^{\tilde{\phi}}_{\alpha,t-x}(i' - 1, j')$ for some state (i', j'). Note that this leaves Process 1 with one more customer that may abandon from station 1 than Process 2. From this point on $\tilde{\phi}$ uses exactly the same allocation decision as ϕ until one of three events occurs; the end of the horizon, an extra abandonment in Process 1 (not seen by Process 2), or Process 2 empties station 1 and ϕ calls for Process 1 to work there. If either of the first two events occurs, the remaining difference in rewards is zero, and Process 2 has received a higher reward than Process 1. That is, ϕ cannot be optimal. If the third event occurs, $\tilde{\phi}$ idles the server until the two processes couple (by abandonment or service completion) or ϕ moves the server to station 2. If there is an extra service seen by Process 1, it receives an extra reward (R_1), and the total rewards coincide (modulo the discounting). Since in each case, the rewards under $\tilde{\phi}$ are higher than that under ϕ, the result follows. □

Since Proposition 2.2 holds for any t, the fact that we can restrict attention to nonidling policies under any of the criteria holds trivially. In the remainder of the paper, we consider only this class of policies.

Remark 2.3 It should be noted that Proposition 2.2 presupposes the existence of an optimal policy for the finite-horizon problem. It is a simple task to show that this is the case (for any fixed t) by applying the results of Theorem 3.1 of [18] with w as defined in Lemma 6.1 below. In the interest of brevity, we have omitted the details for the finite-horizon case. The infinite-horizon cases are included in the Appendix.

2.3 The optimality equations

Let $d(i, j) := \lambda_1 + \lambda_2 + \mu 1_{\{(i,j)\neq(0,0)\}} + i\beta_1 + j\beta_2$. The rate at which transitions occur when the system is in state (i, j) and the server is working on a customer is $d(i, j)$. Since $d(i, j)$ is unbounded in the state space, the decision problem defined by either the rewards or holding cost models is not *uniformizable*. In short, this implies that there is not the typical discrete-time equivalent to the continuous-time problem posed. For a real-valued function f on $\mathbb{X}$, define the following mappings:

$$\mathcal{R}f(i, j) = \lambda_1 f(i + 1, j) + \lambda_2 f(i, j + 1) + i\beta_1 f(i - 1, j) + j\beta_2 f(i, j - 1)$$
$$+ \begin{cases} \mu \max\{R_1 + f(i - 1, j), R_2 + f(i, j - 1)\} & i, j \geq 1, \\ \mu[R_1 + f(i - 1, j)] & i \geq 1, j = 0, \\ \mu[R_2 + f(i, j - 1)] & j \geq 1, i = 0, \\ 0 & (i, j) = (0, 0), \end{cases}$$

and

$$
\begin{aligned}
\mathcal{H}f(i,j) &= i(h_1+\beta_1 P_1) + j(h_2+\beta_2 P_2) + \lambda_1 f(i+1,j) + \lambda_2 f(i,j+1) \\
&\quad + i\beta_1 f(i-1,j) + j\beta_2 f(i,j-1) \\
&\quad + \begin{cases} \mu \min\{f(i-1,j), f(i,j-1)\} & i,j \geq 1, \\ \mu f(i-1,j) & i \geq 1, j=0, \\ \mu f(i,j-1) & j \geq 1, i=0, \\ 0 & (i,j)=(0,0). \end{cases}
\end{aligned}
$$

In each case, the α-discounted reward (resp. cost) optimality equations are defined as $(\alpha + d(i,j))u_\alpha(i,j) = \mathcal{O}u_\alpha(i,j)$, where $\mathcal{O} = \mathcal{R}$ (resp. $\mathcal{H}$). We refer to these equations as the DROE or the DCOE depending on the problem under consideration. Similarly, the average reward or cost optimality equations (AROE or ACOE) are defined by $d(i,j)u(i,j) + g = \mathcal{O}u(i,j)$, where $\mathcal{O} = \mathcal{R}$ (resp. $\mathcal{H}$). The function $u(i,j)$ is called a *relative value function*, and g is the optimal average cost. The next two results state that in each problem and under each criterion, the optimality equations have a solution. The proofs can be found in the Appendix.

Theorem 2.4 *Suppose $\alpha > \max\{\beta_1, \beta_2\}$ and let $\mathcal{O}$ represent the mapping $\mathcal{R}$ or $\mathcal{H}$ depending on the reward or holding cost model. The following hold*:

1. *There exists deterministic policies $\{f_n, n \geq 0\}$ obtaining the maximum/minimum in $(\alpha + d(i,j))u_{n+1,\alpha} := \mathcal{O}u_{n,\alpha}$ (where $u_{0,\alpha} = 0$).*
2. *The function $u_\alpha^* := \lim_{n\to\infty} u_{n,\alpha}$ is a solution of the discounted reward/cost optimality equations and $u_\alpha^* = v_\alpha$.*
3. *There exist deterministic stationary policies f_α^* attaining the maximum/minimum in the discounted reward/cost optimality equations.*

Theorem 2.5 *Let $\mathcal{O}$ represent the mapping $\mathcal{R}$ or $\mathcal{H}$ depending on the reward or holding cost model. The following hold*:

1. *There exists a solution (g^*, u) of the average reward/cost optimality equations. Moreover, g^* is equal to the optimal expected average reward, ρ^*, and u is unique up to additive constants. That is, $g^* = \rho^*(x)$ for all $x \in \mathbb{X}$.*
2. *A deterministic stationary policy is average reward/cost optimal if and only if it achieves the maximum/minimum in the average reward/cost optimality equations.*

The results of Theorems 2.4 and 2.5 imply, for example, that in the discounted reward model, it is optimal to serve at station 1 if $R_1 - R_2 + u_\alpha(i-1,j) - u_\alpha(i,j-1) \geq 0$, while in the holding cost model it is optimal to serve station 1 when $u_\alpha(i-1,j) \leq u_\alpha(i,j-1)$. There is the obvious analogue in the average case. Just as in the discrete-time case, a solution to the average reward/cost optimality equations (g,u) is such that g is the optimal average reward/cost and u is called a *relative value function*. The difference $u(x) - u(y)$ represents the difference in total reward earned by an optimal policy that starts in states x and y, respectively. In the next several sections we discuss when it is optimal to prioritize class 1 or 2 whenever possible.

3 Optimal control

As mentioned in the previous section, the optimality equations (discounted or average rewards or costs) can be used to obtain the structure of an optimal policy by comparing the values (or relative values) when the system starts in different states. In problems that are uniformizable (where $d(i,j)$ can be replaced with a constant), the usual method for doing this comparison is to compare these values term by term. Then using induction through the recursion $(\alpha+d(i,j))u_{n+1,\alpha} := \mathcal{O}u_{n,\alpha}$ inequalities like those above are proved by taking limits. In the current study, we would like to compare states $(i-1,j)$ to $(i,j-1)$. In general, since $d(i-1,j) \neq d(i,j-1)$, the induction is much more difficult (and not doable by these authors); except of course in the case that $d(i-1,j) = d(i,j-1)$ for all $i,j \geq 1$, that is, when $\beta_1 = \beta_2$. This case is considered in the following proposition for a general (nonnegative) cost rate function.

Proposition 3.1 *Suppose $\beta = \beta_1 = \beta_2$ and let $c((i,j),k)$ denote the cost rate in state (i,j) when serving in station $k = 1,2$. Assume that $c(\cdot,a)$ is such that Assumptions A, B, C and Lemma 6.2 (in the Appendix) hold (so that the results of Theorems 2.4 and 2.5 hold). If the following hold*

1. $c((i-1,j),1) \leq c((i,j-1),k)$ *for* $i,j \geq 1$ *and* $k = 1,2$, *and*
2. $c((0,j),2) \leq c((1,j-1),1)$,

then

1. $c((i-1,j),1) + \mu u_\alpha(i-1,j) \leq c((i,j-1),2) + \mu u_\alpha(i,j-1)$ *for all* $i,j \geq 1$, *and*
2. *under either the infinite-horizon discounted cost or average cost criteria, it is optimal to serve at station* 1 *except to avoid unforced idling.*

Proof We show that $u_{n,\alpha}(i-1,j) \leq u_{n,\alpha}(i,j-1)$ for all $i,j \geq 1$ and $n \geq 0$. This, combined with the assumption that $c((i-1,j),1) \leq c((i,j-1),2)$, yields the results upon taking limits. Clearly, this inequality holds for $n = 0$. Assume that it holds for n (which implies that it is optimal to serve at station 1 at epoch $n+1$). Consider $n+1$. The optimality equations $(\alpha + d(i,j))u_{n+1,\alpha} := \mathcal{H}u_{n,\alpha}$ take the form (for $i \geq 2$ and $j \geq 1$)

$$\begin{aligned}&\big(\alpha + d(i-1,j)\big)u_{n+1,\alpha}(i-1,j)\\&\quad= \lambda_1 u_{n,\alpha}(i,j) + \lambda_2 u_{n,\alpha}(i-1,j+1) + (i-1)\beta u_{n,\alpha}(i-2,j)\\&\qquad+ j\beta u_{n,\alpha}(i-1,j-1) + c\big((i-1,j),1\big) + \mu u_{n,\alpha}(i-2,j),\end{aligned}$$

while for $i,j \geq 1$,

$$\begin{aligned}&\big(\alpha + d(i,j-1)\big)u_{n+1,\alpha}(i,j-1)\\&\quad= \lambda_1 u_{n,\alpha}(i+1,j-1) + \lambda_2 u_{n,\alpha}(i,j) + i\beta u_{n,\alpha}(i-1,j-1)\\&\qquad+ (j-1)\beta u_{n,\alpha}(i,j-2) + c\big((i,j-1),1\big) + \mu u_{n,\alpha}(i-1,j-1).\end{aligned}$$

Since $d(i-1,j)=d(i,j-1)$, taking differences and combining like coefficients yields the first statement (with $u_{n,\alpha}$ replacing u_α) via the inductive hypothesis except possibly when considering terms associated with abandonments. Consider only those terms and note

$$\begin{aligned}
&(i-1)\beta u_{n,\alpha}(i-2,j)+j\beta u_{n,\alpha}(i-1,j-1)\\
&\qquad-\big[i\beta u_{n,\alpha}(i-1,j-1)+(j-1)\beta u_{n,\alpha}(i,j-2)\big]\\
&\quad=(i-1)\beta\big[u_{n,\alpha}(i-2,j)-u_{n,\alpha}(i-1,j-1)\big]\\
&\qquad+(j-1)\beta\big[u_{n,\alpha}(i-1,j-1)-u_{n,\alpha}(i,j-2)\big]\\
&\quad\le 0,
\end{aligned}$$

where the inequality holds by applying the inductive hypothesis twice. Now suppose that $i=1$ and $j\ge 1$. The nonidling assumption yields

$$\begin{aligned}
&\big(\alpha+d(i-1,j)\big)u_{n+1,\alpha}(i-1,j)\\
&\quad=\lambda_1 u_{n,\alpha}(i,j)+\lambda_2 u_{n,\alpha}(i-1,j+1)\\
&\qquad+j\beta u_{n,\alpha}(i-1,j-1)+c\big((i-1,j),2\big)+\mu u_{n,\alpha}(i-1,j-1).
\end{aligned}$$

The expressions with coefficient μ cancel, and the induction hypothesis holds for those related to arrivals. The last assumption on the cost function yields the result except possibly with respect to the expressions related to abandonments. However,

$$\begin{aligned}
&j\beta u_{n,\alpha}(i-1,j-1)-\big[\beta u_{n,\alpha}(i-1,j-1)+(j-1)\beta u_{n,\alpha}(i,j-2)\big]\\
&\quad=(j-1)\beta\big[u_{n,\alpha}(i-1,j-1)+u_{n,\alpha}(i,j-2)\big]\\
&\quad\le 0,
\end{aligned}$$

where again the inequality holds by the inductive hypothesis. In each case, the assumptions on the cost function yield $c((i-1,j),1)+\mu u_{n,\alpha}(i-1,j)\le c((i,j-1),2)+\mu u_{n,\alpha}(i,j-1)$ for all n and all $i,j\ge 1$. Taking limits as $n\to\infty$ yields the first result. The second result now holds for the discounted cost case by applying the DCOE. Following the proof of Theorem 4.1 of [12], there exists a subsequence $\{\alpha(n),n\ge 0\}$ such that $u_{\alpha(n)}(i,j)-u_{\alpha(n)}(0,0)\to u(i,j)$, where u satisfies the average cost optimality equations. That is to say that there exists an optimal policy that prioritizes station 1 under either optimality criterion as desired. □

A few notes should be made about the hypotheses of Proposition 3.1. First, in the holding cost model presented, the conditions on the rate functions in Proposition 3.1 translate to precisely what would be expected. That is, $c((i-1,j),1)=(i-1)\times(h_1+\beta P_1)+j(h_2+\beta P_2)\le c((i,j-1),k)=i(h_1+\beta P_1)+(j-1)(h_2+\beta P_2)$ holds if $h_1+\beta P_1\ge h_2+\beta P_2$. On the other hand, in the rewards model, the inequality $R_1\ge R_2$ is implied by $c((i-1,j),1)=-\mu R_1\le c((i,j-1),2)=-\mu R_2$

for $i \geq 2$ (remember c is for costs), but the inequality is $c((0, j), 2) = -\mu R_2 \leq c((1, j-1), 1) = -\mu R_1$ would mean $R_2 \geq R_1$. In short, the results only hold for the case with $R_1 = R_2$. The $R_1 > R_2$ case is covered in what follows, as is the more general holding cost model (without the assumption that $\beta_1 = \beta_2$). Finally, we note that symmetric results hold that yield station 2 should be prioritized. We believe not only are the next set of results of interest, but also the methodologies may be of use for a wide range of related problems.

3.1 The rewards model

In this section we provide conditions under which a *priority rule* holds in the reward model. Originally, one might conjecture that $R_1 \geq R_2$ is sufficient to guarantee the optimality of a rule that prioritizes station 1. The following (counter)example shows that this is not always the case.

Example 3.2 Suppose that we have the following model inputs: $\lambda_1 = 0.1$; $\lambda_2 = 0.1$; $\mu = 1$; $\beta_1 = 0.1$; $\beta_2 = 3$; $R_1 = 2.0$; $R_2 = 1.0$. With these inputs, the average reward of a policy that serves at station 1 (except to avoid idling) is $\rho_1 = 0.002809$, while the optimal policy has average reward $\rho^* = 0.003185$, a 13.4% increase.

Figure 1 depicts the optimal policy for this example. Not only is it not strictly a priority policy, but since it is nonidling, it is also nonmonotone in the number of customers in station 1.

The following provides conditions under which it is optimal to always serve at one station or the other (except to avoid unforced idling) and is the main result of the section.

Theorem 3.3 *Suppose* $\beta_1 \geq \beta_2$ *and* $R_1 \geq R_2$. *Then* $R_1 - R_2 + u(i-1, j) - u(i, j-1) \geq 0$ *for all* $i, j \geq 1$, *and an optimal policy exists that always serves at station* 1, *except to avoid unforced idling. By symmetry, if* $\beta_2 \geq \beta_1$ *and* $R_2 \geq R_1$, *then*

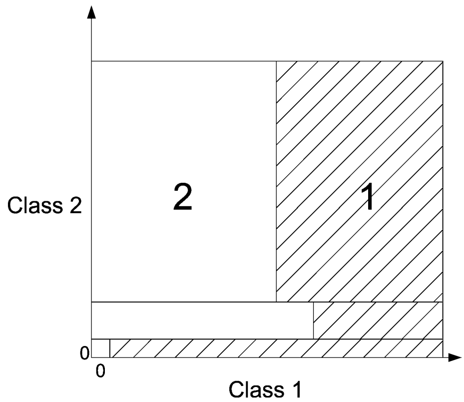

Fig. 1 Graphical depiction of the optimal policy for Example 3.2

$R_1 - R_2 + u(i-1, j) - u(i, j-1) \le 0$ for all $i, j \ge 1$, and an optimal policy exists that always serves at station 2, except to avoid unforced idling.

The proof of Theorem 3.3 is delayed until we have proved the next proposition. Before proceeding however, consider again Example 1. Note that in the case that $R_1 \ge R_2$ and $\beta_2 > \beta_1$, the decision-maker has two competing objectives. First, there is a desire to maximize rewards, and so station 1 should be prioritized. On the other hand, if β_2 is too high, all of the station 2 customers may abandon while the server is clearing station 1; resulting in future server idleness and corresponding lost rewards. So a balance must be struck between maximizing rewards and avoiding idleness. Both are achieved by serving at station 1 when $\beta_1 \ge \beta_2$.

Proposition 3.4 *The following hold for any fixed t:*

1. *Suppose $\beta_1 \ge \beta_2$ and $R_1 \ge R_2$. Then $R_1 - R_2 + u_{\alpha,t}(i-1, j) - u_{\alpha,t}(i, j-1) \ge 0$ for all $i, j \ge 1$.*
2. *Suppose $\beta_2 \ge \beta_1$ and $R_2 \ge R_1$. Then $R_1 - R_2 + u_{\alpha,t}(i-1, j) - u_{\alpha,t}(i, j-1) \le 0$ for all $i, j \ge 1$.*

Proof To prove the first result, fix t and consider $u_{\alpha,t}(i-1, j) - u_{\alpha,t}(i, j-1)$. Define two processes on the same probability space. Process 1 starts in state $(i-1, j)$ and serves in the same station as process 2, whenever possible. Process 2 starts in state $(i, j-1)$ and uses an optimal policy. Since both processes are defined on the same space, we assume that they see the same arrivals and potential services. If an arrival is the first event at time t_0 say, the relative position of the two processes remains the same, and they each enter new states. There are now $t - t_0$ time units remaining. We relabel the new states as the initial states and continue with the same argument that follows.

As for the abandonments, assume that we generate the first $i-1$ and the first $j-1$ customers in each queue so that both processes see the same abandonments. If any of these events occur first, again, the relative positions of each process remain the same, and we continue as before. For the remaining customer (an extra at station 1 in process 2 and an extra at station 2 in process 1), we generate a single exponential with rate β_1. If this event occurs first, then both processes see an extra abandonment with probability $\frac{\beta_2}{\beta_1}$. This implies that the difference in the remaining rewards is $u_{\alpha,t-t_0}(i-1, j-1) - u_{\alpha,t-t_0}(i-1, j-1) = 0$. With probability $\frac{\beta_1-\beta_2}{\beta_1}$ it generates an abandonment in station 1 for process 2 (not seen by process 1). The remaining rewards are $u_{\alpha,t-t_0}(i-1, j) - u_{\alpha,t-t_0}(i-1, j-1) \ge 0$, where the inequality is due to the first result of Proposition 2.1. The assumption that $R_1 \ge R_2$ yields the first inequality in this case.

Consider now the services. Recall that process 2 uses the optimal policy. Assume that process 1 serves in the same station as process 2, whenever possible. Since each service can be constructed so that both processes see the same service times, the relative position of each process remains the same except in the case of $i - 1 = 0$ and process 2 serves at station 1. Suppose that this is the case. At this time the potentially suboptimal policy for process 1 serves at station 2. If the service is the next event, the instantaneous rewards are different, and the difference in the remaining rewards is

$$e^{-\alpha(t-t_0)}(R_2 - R_1) + u_{\alpha,t-t_0}(0, j-1) - u_{\alpha,t-t_0}(0, j-1) = e^{-\alpha(t-t_0)}(R_2 - R_1).$$

Adding $R_1 - R_2$ yields $(R_1 - R_2)(1 - e^{-\alpha(t-t_0)}) \geq 0$, as desired.

Consider again $u_{\alpha,t}(i-1, j) - u_{\alpha,t}(i, j-1)$ for generic $i, j \geq 1$. Let p be the probability that the processes enter states $(0, j')$ and $(1, j'-1)$ for some j'. The previous arguments imply that $u_{\alpha,t}(i-1, j) - u_{\alpha,t}(i, j-1) \geq p(R_2 - R_1) \geq R_2 - R_1$, where the inequality follows since $R_1 \geq R_2$. The result is proven. The remaining result holds by symmetry. □

Since t was arbitrary, by taking limits as $t \to \infty$ Theorem 3.3 is immediate.

One might note that the proof of Proposition 3.4 relies on two important facts. First that no reward or costs are accrued between events and second that the instantaneous rewards or costs are not state dependent. Neither of these hold for the holding cost model which is considered in the next section.

3.2 Holding costs

As an alternative to the methods of the previous section, the classic "c–μ" result was shown using an *interchange* argument (cf. Varaiya and Buyokkoc [7] or Nain [17]). In essence, an index for each station is created (the holding cost times the service rate). The station with the highest index receives the highest priority. The argument is that any policy that violates this priority rule can be improved by rearranging the order in which customers are served in accordance with the index. Two processes are defined on the same space that use the various policies. Since all customers that arrive at a particular station will be served and served in the order in which they arrived, the two processes can be made to couple. The process that follows the index rule drains cost earlier and therefore minimizes the total cost. The difficulty in the current study is in the assumption that the two processes can be made to couple. Indeed, some customers may abandon awaiting service in one process while they have their service completed in the other. If this happens, there is no way to guarantee the processes will couple. In what follows, we discuss the holding cost model and what can be done to alleviate this difficulty. The main results of this section are captured in the following theorem. Its proof is divided into several steps.

Theorem 3.5 *Suppose the following hold*:

1. $h_1 + \beta_1 P_1 \geq (\leq)\, h_2 + \beta_2 P_2$,
2. $\beta_2 \geq (\leq)\, \beta_1$.

Then under either the infinite-horizon discounted cost or average cost criteria, *there exists an optimal policy that prioritizes station* 1 (2) *except to avoid unforced idling*.

Our original intuition was that $h_1 + \beta_1 P_1 \geq h_2 + \beta_2 P_2$ should be sufficient to prioritize station 1. After all, this would be in line with classic results. The next example addresses the question of necessity and sufficiency of the added inequality $\beta_2 \geq \beta_1$.

Example 3.6 Suppose $\lambda_1 = 2; \lambda_2 = 2.5; \mu = 3; \beta_1 = 0.9; \beta_2 = 1; h_1 = 1.5; h_2 = 1; P_1 = 1; P_2 = 0.5$.

Note that $h_1 + \beta_1 P_1 = 2.4 \geq 1.5 = h_2 + \beta_2 P_2$. The optimal policy (computed via *Matlab*) is to work at station 1 unless there are no customers at station 1. This same policy is optimal if we let $\beta_1 = 1.1 > \beta_2$. That is, the hypotheses of Theorem 3.5 are sufficient but not necessary. If we let $\beta_1 = 2$, then $h_1 + \beta_1 P_1 = 3.5 \geq 1.5 = h_2 + \beta_2 P_2$. However, the optimal policy is to serve at station 2; following our intuition could lead to using a priority rule that is exactly the opposite of what is optimal!

As has been alluded to, the classic methods of a sample path argument or interchange argument cannot be applied directly. We have also mentioned that the problem is not uniformizable so that there is not a discrete-time equivalent Markov decision process. One might suggest that we could truncate the state space, making it uniformizable, prove the results on the truncated space, and take limits as the truncation level approaches infinity. This approach is also suggested by Assumption A in the Appendix. The next example shows that care must be taken when choosing the truncation. Suppose that each queue is truncated when it reaches $L = 20$; excess customers are lost.

Example 3.7 Let $\lambda_1 = 2; \lambda_2 = 2.5; h_1 = 1.01, h_2 = 1.0; \mu_1 = \mu_2 = 4.6;\ \beta_1 = \beta_2 = 0$.

Note that Example 3.7 does not include abandonments. The optimal policy for the example is depicted in Fig. 2. Close to the boundary, it may not be optimal to prioritize station 1 in spite of the fact that $h_1 \geq h_2$. In the original untruncated model, each customer that arrives to station k increases the cost per unit time by h_k, $k = 1, 2$. In the truncated system, when the number of customers in station 1 is 20, a customer arriving at station 1 does not increase the cost, while a station 2 arrival (as long as station 2 has less than 20 customers) increases the cost by h_2; it may be advantageous to keep station 1 full.

To this end, we consider the following *equivalent* formulation. Suppose that the state space is $\mathcal{Y} = \{(I, i) : 0 \leq i \leq I < \infty\}$, where I represents the current number of

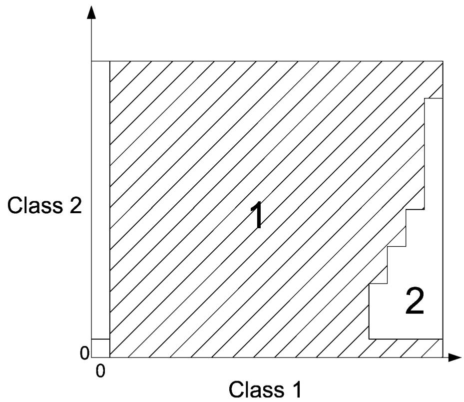

Fig. 2 Graphical depiction of the optimal policy for Example 3.7

customers in the system, and i is the number at station 1. Replacing j with $(I-i)$, we note that $d(i, I-i) = m(I,i) := \lambda_1 + \lambda_2 + \mu + i\beta_1 + (I-i)\beta_2$. The DCOE are now (there is also the obvious analogue for the ACOE)

$$\begin{aligned}\big(\alpha + m(I,i)\big)u_\alpha(I,i) = {} & i(h_1 + \beta_1 P_1) + (I-i)(h_2 + \beta_2 P_2) + \lambda_1 u_\alpha(I+1, i+1) \\ & + \lambda_2 u_\alpha(I+1, i) + \mu \min\big\{u_\alpha(I-1, i-1), u_\alpha(I-1, i)\big\} \\ & + i\beta_1 u_\alpha(I-1, i-1) + (I-i)\beta_2 u_\alpha(I-1, i). \end{aligned} \tag{3.1}$$

3.2.1 Finite state approximation

Recall that a uniformizable continuous-time MDP has an *equivalent* discrete-time formulation where the optimal policies coincide, and the optimal values are within a multiplicative constant of each other (see [16] or [20]). Suppose that the maximum number of customers allowed in the system at any time is L, where L is finite. Let $\overline{\beta} = \max\{\beta_1, \beta_2\}$. Thus, the abandonment rate from the system is bounded above by $L\overline{\beta}$. Since under these assumptions the Markov decision process is uniformizable, let $\Psi_L := \lambda_1 + \lambda_2 + \mu + L\overline{\beta} = 1$, where the last equality is without loss of generality. Since in this section L will be fixed, we suppress dependence on L. For example, the uniformized discount factor $\delta_L = \frac{\Psi_L}{\alpha + \Psi_L}$ will simply be denoted δ.

It remains to describe what happens when a customer arrives at a station when there are already L total customers in the system. When $I = L$, a customer arriving at station 2 is lost forever. When $I = L$, $i < L$, and a customer arrives at station 1, a customer is *removed* from station 2 (without penalty), and the arriving customer joins the queue at station 1. When $i = L$, any arriving customer is lost. That is, when an arrival occurs to station 1 in state (L, i), the next state is $(L, (i+1) \wedge L)$. We have already discussed after Example 3.7 the difficulty in truncating the queue lengths at each station. The dynamics on the boundary alleviate that concern by making station 1 arrivals increase the cost while actually decreasing the cost at station 2. Since this is only a change on the boundary, when we take limits (as the boundary moves off to infinity), we still approach the original problem. The discrete-time optimality criteria are defined for a fixed policy π by

$$v^\pi_{N,\delta}(x) := \mathbb{E}^\pi_x \sum_{n=0}^{N-1} \big[\delta^n C\big(X_n, d_n(X_n)\big)\big], \tag{3.2}$$

$$v^\pi_\delta := \lim_{N\to\infty} v^\pi_{N,\delta}(x), \tag{3.3}$$

where $\{X_n, n \geq 0\}$ denotes the stochastic process representing the state at decision epoch n. Equations (3.2) and (3.3) define the N-stage expected discounted cost and the infinite-horizon expected discounted cost, respectively. Again, in each case, we define the optimal values $y(i,j) := \inf_{\pi \in \Pi} y^\pi(i,j)$, where $y = v_{N,\delta}$ or v_δ depending on the optimality criterion.

The (discrete-time) finite-horizon optimality equations for $1 \le i < I < L$ are ($v_{0,\delta} = 0$)

$$\begin{aligned} v_{n+1,\delta}(I,i) = {} & i(h_1 + \beta_1 P_1) + (I-i)(h_2 + \beta_2 P_2) \\ & + \delta\big(\lambda_1 v_{n,\delta}(I+1,i+1) + \lambda_2 v_{n,\delta}(I+1,i) \\ & + \mu \min\big\{v_{n,\delta}(I-1,i-1), v_{n,\delta}(I-1,i)\big\} \\ & + \big[L\overline{\beta} - i\beta_1 - (I-i)\beta_2\big] v_{n,\delta}(I,i) + i\beta_1 v_{n,\delta}(I-1,i-1) \\ & + (I-i)\beta_2 v_{n,\delta}(I-1,i)\big). \end{aligned} \tag{3.4}$$

When $1 \le i = I < L$,

$$\begin{aligned} v_{n+1,\delta}(I,I) = {} & I(h_1 + \beta_1 P_1) + \delta\big(\lambda_1 v_{n,\delta}(I+1,I+1) + \lambda_2 v_{n,\delta}(I+1,I) \\ & + (\mu + I\beta_1) v_{n,\delta}(I-1,I-1) + [L\overline{\beta} - I\beta_1] v_{n,\delta}(I,I)\big). \end{aligned}$$

For $i = 0$ and $I < L$,

$$\begin{aligned} v_{n+1,\delta}(I,0) = {} & I(h_2 + \beta_2 P_2) + \delta\big(\lambda_1 v_{n,\delta}(I+1,1) + \lambda_2 v_{n,\delta}(I+1,0) \\ & + \mu v_{n,\delta}(I-1,0) + [L\overline{\beta} - I\beta_2] v_{n,\delta}(I,0) + I\beta_2 v_{n,\delta}(I-1,0)\big). \end{aligned}$$

When $I = L$ and $i \ge 1$,

$$\begin{aligned} v_{n+1,\delta}(L,i) = {} & i(h_1 + \beta_1 P_1) + (L-i)(h_2 + \beta_2 P_2) + \delta\big(\lambda_1 v_{n,\delta}\big(L,(i+1)\wedge L\big) \\ & + \lambda_2 v_{n,\delta}(L,i) + \mu \min\big\{v_{n,\delta}(L-1,i-1), v_{n,\delta}(L-1,i)\big\} \\ & + \big[L\overline{\beta} - i\beta_1 - (L-i)\beta_2\big] v_{n,\delta}(L,i) + i\beta_1 v_{n,\delta}(L-1,i-1) \\ & + (L-i)\beta_2 v_{n,\delta}(L-1,i)\big), \end{aligned}$$

and for $i = 0$ and $I = L$,

$$\begin{aligned} v_{n+1,\delta}(L,0) = {} & L(h_2 + \beta_2 P_2) + \delta\big(\lambda_1 v_{n,\delta}(L,1) + \lambda_2 v_{n,\delta}(L,0) + \mu v_{n,\delta}(L-1,0) \\ & + [L\overline{\beta} - L\beta_2] v_{n,\delta}(L,0) + L\beta_2 v_{n,\delta}(L-1,0)\big). \end{aligned}$$

Note that it is optimal to serve customers at station 1 in state (I,i) when $v_{n,\delta}(I-1,i-1) \le v_{n,\delta}(I-1,i)$. The discrete-time discounted cost optimality equations are precisely the same with $v_{n+1,\delta}$ and $v_{n,\delta}$ replaced with v_δ. In each case, it is well known that the optimal values satisfy the optimality equations (cf. Chap. 6 of [19]). For fixed I, let $\Delta_2 v_{n,\delta}(I,i) = v_{n,\delta}(I,i+1) - v_{n,\delta}(I,i)$. Thus, in state $(I+1,i+1)$, it is optimal to serve at station 1 if $\Delta_2 v_{n,\delta}(I,i) \ge 0$. For $I < L$, let $\Delta_1 v_{n,\delta}(I,i) = v_{n,\delta}(I+1,i) - v_{n,\delta}(I,i)$.

Proposition 3.8 *Suppose the following hold*:

1. $h_1 + \beta_1 P_1 \geq (\leq) h_2 + \beta_2 P_2$,
2. $\beta_2 \geq (\leq) \beta_1$.

Then

1. $\Delta_1 v_{n,\delta}(I,i) \geq (\leq) 0$ *for all* $i \leq I < L$ *and* $n \geq 0$,
2. $\Delta_2 v_{n,\delta}(I,i) \geq (\leq) 0$ *for all* $i \leq I \leq L$ *and for all* $n \geq 0$,
3. *The previous inequalities hold when* $v_{n,\delta}$ *is replaced by* v_δ.

Proof We prove the result in the "$\geq$" direction; the opposite direction holds by symmetry. To ease notation, assume that $\delta = 1$; the case for $\delta < 1$ is analogous. The fact that both inequalities hold when $n = 0$ is trivial. Assume that they hold for n and consider $n + 1$. The second inductive hypothesis implies that it is optimal to serve at station 1 at time $n + 1$ except to avoid idling. Suppose $I = L - 1$. If $i = L - 1$, then note that an arrival at station 1 in states $(L, L - 1)$ or $(L - 1, L - 1)$ leads to the next state being (L, L). Similarly, an arrival at station 2 in either of those same states leads to $(L, L - 1)$. Thus,

$$\begin{aligned}\Delta_1 v_{n+1,\delta}(L-1, L-1) = h_2 + \beta_2 P_2 &+ \mu \Delta_1 v_{n,\delta}(L-2, L-2) \\ &+ \big[L\overline{\beta} - (L-1)\beta_1 - \beta_2\big]\Delta_1 v_{n,\delta}(L-1, L-1) \\ &+ (L-1)\beta_1 \Delta_1 v_{n,\delta}(L-2, L-2).\end{aligned}$$

The inductive hypothesis yields the result in each case. Similarly, if $i = 0$ (station 2 arrivals in $(L, 0)$ or $(L - 1, 0)$ both lead to $(L, 0)$),

$$\begin{aligned}&\Delta_1 v_{n+1,\delta}(L-1, 0) \\ &\quad = h_2 + \beta_2 P_2 + \lambda_1 \Delta_1 v_{n,\delta}(L, 1) + \big(\mu + (L-1)\beta_2\big)\Delta_1 v_{n,\delta}(L-2, 0) \\ &\qquad + [L\overline{\beta} - L\beta_2]\Delta_1 v_{n,\delta}(L-1, 0).\end{aligned}$$

For $0 < i < L - 1$,

$$\begin{aligned}&\Delta_1 v_{n+1,\delta}(L-1, i) \\ &\quad = h_2 + \beta_2 P_2 + \lambda_1 \Delta_1 v_{n,\delta}(L-1, i+1) + \mu \Delta_1 v_{n,\delta}(L-2, i-1) \\ &\qquad + \big[L\overline{\beta} - i\beta_1 - (L-i)\beta_2\big]\Delta_1 v_{n,\delta}(L-1, i) \\ &\qquad + i\beta_1 \Delta_1 v_{n,\delta}(L-2, i-1) + (L-1-i)\beta_2 \Delta_1 v_{n,\delta}(L-2, i),\end{aligned}$$

and the inductive hypothesis yields the result. Next, consider $I < L - 1$ and $i = I$. We have

$$\begin{aligned}&\Delta_1 v_{n+1,\delta}(I, I) \\ &\quad = h_2 + \beta_2 P_2 + \lambda_1 \Delta_1 v_{n,\delta}(I+1, I+1) + \lambda_2 \Delta_1 v_{n,\delta}(I+1, I) \\ &\qquad + (\mu + I\beta_1)\Delta_1 v_{n,\delta}(I-1, I-1) + [L\overline{\beta} - I\beta_1 - \beta_2]\Delta_1 v_{n,\delta}(I, I).\end{aligned}$$

The inductive hypotheses yield the result. For $i=0$,

$$\begin{aligned}\Delta_1 v_{n+1,\delta}(I,0) = h_2 + \beta_2 P_2 + \lambda_1 \Delta_1 v_{n,\delta}(I+1,1) + \lambda_2 \Delta_1 v_{n,\delta}(I+1,0)\\ + \mu \Delta_1 v_{n,\delta}(I-1,0) + \big[L\overline{\beta} - (I+1)\beta_2\big]\Delta_1 v_{n,\delta}(I,0)\\ + I\beta_2 \Delta_1 v_{n,\delta}(I-1,0).\end{aligned}$$

For $0<i<I$,

$$\begin{aligned}\Delta_1 v_{n+1,\delta}(I,i) = h_2 + \beta_2 P_2 + \lambda_1 \Delta_1 v_{n,\delta}(I+1,i+1) + \lambda_2 \Delta_1 v_{n,\delta}(I+1,i)\\ + \mu \Delta_1 v_{n,\delta}(I-1,i-1) + \big[L\overline{\beta} - i\beta_1 - (I+1-i)\beta_2\big]\Delta_1 v_{n,\delta}(I,i)\\ + i\beta_1 \Delta_1 v_{n,\delta}(I-1,i-1) + (I-i)\beta_2 \Delta_1 v_{n,\delta}(I-1,i).\end{aligned}$$

In each case, the inductive hypothesis yields the result. For $I=0$ (so that $i=0$), we have

$$\begin{aligned}\Delta_1 v_{n+1,\delta}(0,0) = h_2 + \beta_2 P_2 + \lambda_1 \Delta_1 v_{n,\delta}(1,1) + \lambda_2 \Delta_1 v_{n,\delta}(1,0)\\ + [L\overline{\beta} - \beta_2]\Delta_1 v_{n,\delta}(0,0).\end{aligned}$$

This completes the proof of the first inequality. To prove the second inequality, consider first $I=L$ and $0<i<L$. If $i=L-1$,

$$\begin{aligned}&\Delta_2 v_{n+1,\delta}(L,L-1)\\ &\quad= h_1 + \beta_1 P_1 - [h_2 + \beta_2 P_2] + \lambda_1\big[v_{n,\delta}(L,L) - v_{n,\delta}(L,L)\big]\\ &\qquad + \lambda_2\big[\Delta_2 v_{n,\delta}(L,L-1)\big] + \mu\big[\Delta_2 v_{n,\delta}(L-1,L-2)\big] + [L\overline{\beta} - L\beta_1]v_{n,\delta}(L,L)\\ &\qquad - \big[L\overline{\beta} - (L-1)\beta_1 - \beta_2\big]v_{n,\delta}(L,L-1) + L\beta_1 v_{n,\delta}(L-1,L-1)\\ &\qquad - (L-1)\beta_1 v_{n,\delta}(L-1,L-2) - \beta_2 v_{n,\delta}(L-1,L-1)\\ &\quad= h_1 + \beta_1 P_1 - [h_2 + \beta_2 P_2] + \lambda_2\big[\Delta_2 v_{n,\delta}(L,L-1)\big]\\ &\qquad + \mu\big[\Delta_2 v_{n,\delta}(L-1,L-2)\big] + [L\overline{\beta} - L\beta_1]\Delta_2 v_{n,\delta}(L,L-1)\\ &\qquad + (L-1)\beta_1 \Delta_2 v_{n,\delta}(L-1,L-2) + (\beta_2 - \beta_1)\Delta_1 v_{n,\delta}(L-1,L-1).\end{aligned}$$

The second inductive hypothesis holds in each case involving $v_{n,\delta}$, save the last one, where the first inductive hypothesis yields the result. Consider now the case where $i=0$. Then

$$\begin{aligned}\Delta_2 v_{n+1,\delta}(L,0) = h_1 + \beta_1 P_1 - [h_2 + \beta_2 P_2] + \lambda_1 \Delta_2 v_{n,\delta}(L,1) + \lambda_2 \Delta_2 v_{n,\delta}(L,0)\\ + \big[L\overline{\beta} - \beta_1 - (L-1)\beta_2\big]\Delta_2 v_{n,\delta}(L,0)\\ + (\beta_2 - \beta_1)\Delta_1 v_{n,\delta}(L-1,0) + (L-1)\beta_2 \Delta_2 v_{n,\delta}(L-1,0).\end{aligned}$$

The same argument as above yields the result. Suppose $I=L$ and $0<i<L-1$. A little algebra yields

$$
\begin{aligned}
&\Delta_2 v_{n+1,\delta}(L,i)\\
&\quad= h_1+\beta_1 P_1-[h_2+\beta_2 P_2]+\lambda_1\Delta_2 v_{n,\delta}(L,i+1)+\lambda_2\Delta_2 v_{n,\delta}(L,i)\\
&\qquad+\mu\Delta_2 v_{n,\delta}(L-1,i-1)+\big[L\overline{\beta}-(i+1)\beta_1-(L-i-1)\beta_2\big]\Delta_2 v_{n,\delta}(L,i)\\
&\qquad+i\beta_1\Delta_2 v_{n,\delta}(L-1,i-1)+(L-i-1)\beta_2\Delta_2 v_{n,\delta}(L-1,i)\\
&\qquad+(\beta_2-\beta_1)\Delta_1 v_{n,\delta}(L-1,i).
\end{aligned}
$$

The same argument as in the previous cases holds. When $I<L$, there are also several cases to consider. However, for $i=I-1$, note

$$
\begin{aligned}
&\Delta_2 v_{n+1,\delta}(I,I-1)\\
&\quad= h_1+\beta_1 P_1-[h_2+\beta_2 P_2]+\lambda_1\Delta_2 v_{n,\delta}(I+1,I)+\lambda_2\Delta_2 v_{n,\delta}(I+1,I-1)\\
&\qquad+\mu\Delta_2 v_{n,\delta}(I-1,I-2)+[L\overline{\beta}-I\beta_1]\Delta_2 v_{n,\delta}(I,I-1)\\
&\qquad+(I-1)\beta_1\Delta_2 v_{n,\delta}(I-1,I-2)+(\beta_2-\beta_1)\Delta_1 v_{n,\delta}(I-1,I-1).
\end{aligned}
$$

The result follows. For $i=0$,

$$
\begin{aligned}
&\Delta_2 v_{n+1,\delta}(I,0)\\
&\quad= h_1+\beta_1 P_1-[h_2+\beta_2 P_2]+\lambda_1\Delta_2 v_{n,\delta}(I+1,1)+\lambda_2\Delta_2 v_{n,\delta}(I+1,0)\\
&\qquad+\big[L\overline{\beta}-\beta_1-(I-1)\beta_2\big]\Delta_2 v_{n,\delta}(I,0)+(\beta_2-\beta_1)\Delta_1 v_{n,\delta}(I,0)\\
&\qquad+(I-1)\Delta_2\beta_2 v_{n,\delta}(I-1,0).
\end{aligned}
$$

For $0<i<I-1$,

$$
\begin{aligned}
&\Delta_2 v_{n+1,\delta}(I,i)\\
&\quad= h_1+\beta_1 P_1-[h_2+\beta_2 P_2]+\lambda_1\Delta_2 v_{n,\delta}(I+1,i+1)+\lambda_2\Delta_2 v_{n,\delta}(I+1,i)\\
&\qquad+\mu\Delta_2 v_{n,\delta}(I-1,i-1)+\big[L\overline{\beta}-(i+1)\beta_1-(I-i-1)\beta_2\big]\Delta_2 v_{n,\delta}(I,i)\\
&\qquad+(\beta_2-\beta_1)\Delta_1 v_{n,\delta}(I,i-1)+i\beta_1\Delta_2 v_{n,\delta}(I-1,i-1)\\
&\qquad+(I-i-1)\beta_2\Delta_2 v_{n,\delta}(I-1,i),
\end{aligned}
$$

which is nonnegative, as desired. The third result follows by noting that $v_{n,\delta}\to v_\delta$. $\square$

3.2.2 *Convergence to the countable state model*

In this section we show that the infinite-horizon discounted cost value function for the truncated system, $v_{\alpha,L}$, converges to that in the original system. We have dispensed with the assumptions that $\delta_L=\Psi_L=1$ and added back in the dependence on L. Note that $v_{\alpha,L}$ is the unique (bounded) vector satisfying the discrete-time infinite-horizon δ_L-discounted cost optimality equations (for $0<i<I<L$). So,

$$\begin{aligned}(\alpha+\Psi_L)v_{\alpha,L}(I,i) &= ih_1+(I-i)h_2+i\beta_1P_1+(I-i)\beta_2P_2\\ &\quad+\lambda_1v_{\alpha,L}(I+1,i+1)+\lambda_2v_{\alpha,L}(I+1,i)\\ &\quad+\mu\min\{v_{\alpha,L}(I-1,i-1),v_{\alpha,L}(I-1,i)\}\\ &\quad+\big[L\overline{\beta}-i\beta_1-(I-i)\beta_2\big]v_{\alpha,L}(I,i)+i\beta_1v_{\alpha,L}(I-1,i-1)\\ &\quad+(I-i)\beta_2v_{\alpha,L}(I-1,i),\end{aligned}\tag{3.5}$$

where the above expression can be obtained by replacing $v_{n,\delta}$ in (3.4) with $v_{\alpha,L}$ and using a little algebra. For completeness, we assume that $v_{\alpha,L}(I,i)=0$ for $I>L$. The next result shows that the limit of $v_{\alpha,L}$ exists.

Lemma 3.9 *$v_{\alpha,L}$ is (pointwise) monotone in L.*

Proof We need to prove that $v_{\alpha,L+1}(I,i)\geq v_{\alpha,L}(I,i)$ for all $0\leq i\leq I$ and all $L\geq 0$. First note that for $I\geq L+1$, the result holds trivially (by assumption). To complete the proof, follow the sample paths of two processes defined on the same probability space and starting in the same state where $I\leq L$. Suppose that π^*_{L+1} is an optimal policy for the state space bounded by $L+1$. Let π_L be a policy that serves at exactly the same station as π^*_{L+1}. Process 1 uses policy π^*_{L+1} and operates on the states such that $I\leq L+1$. Process 2 uses policy π_L and operates on the states such that $I\leq L$. Now since both processes use the same policy when $I<L$, as long as the total number of customers is less than L, they see the same arrivals, services and abandonments, and, therefore the same costs. Consider the first time the processes enter a state with the number of customers equal to L, say (L,i'). If a service or abandonment is the next event, both processes remain coupled until the next time they have L customers in the system. If a class 1 arrival occurs, and $i'\neq L$, both processes see an increase in the number of class 1 customers. Process 1 is in state $(L+1,i'+1)$, while Process 2 is in state $(L,i'+1)$. After this time, Process 2 does not serve at station 2, until there is either an extra abandonment or an extra service at station 2. In particular, if the optimal policy tells Process 1 to serve at station 1, so does Process 2. If it says to work at station 2, Process 2 idles until the service is complete (or an extra abandonment occurs). Thus, since Process 1 accrues costs at a higher rate and is always in a lower state (according to the cost function), we have

$$v_{\alpha,L+1}(I,i)\geq v^{\pi_L}_{\alpha,L}(I,i)\geq v_{\alpha,L}(I,i).$$

Since the initial state was arbitrary, the result follows. □

Lemma 3.9 implies that $v_{\alpha,L}$ converges as L increases. Let $v_{\alpha,\infty}$ denote this (possibly infinite) limit. A little algebra in (3.5) yields, for $I<L$,

$$\begin{aligned}\big(\alpha+m(I,i)\big)v_{\alpha,L}(I,i) &= ih_1+(I-i)h_2+i\beta_1P_1+(I-i)\beta_2P_2\\ &\quad+\lambda_1v_{\alpha,L}(I+1,i+1)+\lambda_2v_{\alpha,L}(I+1,i)\\ &\quad+\mu\min\{v_{\alpha,L}(I-1,i-1),v_{\alpha,L}(I-1,i)\}\\ &\quad+i\beta_1v_{\alpha,L}(I-1,i-1)+(I-i)\beta_2v_{\alpha,L}(I-1,i),\end{aligned}$$

which for $I < L$ is precisely the same as (3.1). Thus, as $L \to \infty$, $v_{\alpha,L} \to v_{\alpha,\infty} = v_\alpha$. This leads to the proof of Theorem 3.5.

Proof of Theorem 3.5 Since $\Delta_1 v_{\alpha,L} \geq 0$ for all L, the first inequality follows from the fact $v_{\alpha,L} \to v_\alpha$. Similarly, define $u_\alpha(I,i) = v_\alpha(I,i) - v_\alpha(0,0)$. Since u_α (and αv_α) converges along a subsequence to a solution of the CTDCOE (see the proof of Theorem 4.1 of [12]), $(\rho^*, u(I,i))$, the inequality holds in the average case as well. The result follows. □

4 Numerical results

In this section we discuss the improvements that may be possible when the abandonment rates are such that the intuitive index policy (either give priority to the largest R_i or to the largest $h_i + \beta_i P_i$) is not guaranteed to be optimal. In both the rewards and holding cost models, we discern under what conditions one should be careful in the choice of policy and also try to show how much system performance may be impacted.

4.1 The rewards model

We provide results for a system with $\lambda_1 = 1$ and $\lambda_2 = \mu = 4$. We initially set $R_1 = 10$ and $R_2 = 5$, to model a system where, in the overall offered demand, there is a small proportion of high-revenue customers. Giving priority to the high-reward customers maximizes short-term rewards, and if the abandonment rates are ordered such that $\beta_1 \geq \beta_2$, then according to Theorem 3.3, this policy is also optimal in the long-run. We are interested in seeing what happens when $\beta_2 > \beta_1$. In this case, one can think that there may be a trade-off between maximizing short-term reward and minimizing the amount of offered demand that is lost through abandonments.

We studied a truncated system with buffer size 20 for both classes. In all of the results that follow, we use ρ_1 to denote the average reward for a policy that gives priority to queue 1, while ρ^* is the average reward for the optimal policy.

First, we fix $\beta_2 = 2.0$ and observe the effect of varying β_1. The results in Table 1 demonstrate that the improvement in using the optimal policy increases as β_1 decreases, as one would expect. (In Tables 1, 2, and 3, the last column indicates the form of the optimal policy. P1 denotes priority to class 1, P2 denotes priority to

Table 1 Rewards model, varying β_1

β_1	ρ_1	ρ^*	% from optimal	Policy
0	0.353	0.394	10.4	P2
0.1	0.336	0.358	6.1	T1
0.2	0.320	0.332	3.6	T1
0.5	0.281	0.281	0	P1
1.0	0.233	0.233	0	P1
2.0	0.172	0.172	0	P1

Table 2 Rewards model, varying β_2

β_2	ρ_1	ρ^*	% from optimal	Policy
1.0	0.585	0.605	3.3	T1
2.0	0.336	0.358	6.1	T1
5.0	0.135	0.147	8.2	T1
10.0	0.632	0.678	6.8	T1

Table 3 Rewards model, varying R_2

R_2	ρ_1	ρ^*	% from optimal	Policy
1	0.208	0.208	0	P1
2	0.240	0.242	0.8	T1
5	0.336	0.358	6.1	T1
9	0.464	0.516	10.1	T1

class 2, and T1 gives priority to class 1 if the number of class 1 customers is greater than a (state-dependent) threshold.) As β_1 decreases, it becomes advantageous to devote more effort to queue 2, to avoid excessive lost demand, as customers are less likely to be lost from queue 1. For β_1 small, the optimal policy actually gives priority to queue 2. At $\beta_1 = 0.5$, even though β_1 is still less than β_2, giving priority to queue 1 becomes optimal.

Equivalently, we would expect the trade-off described above to become more significant as β_2 grows and R_2 approaches R_1. Both of these expectations are confirmed in Tables 2 and 3. Table 2 has β_1 fixed at 0.1 and varies β_2, while Table 3 fixes $\beta_1 = 0.1$, $\beta_2 = 2.0$, and varies R_2 (here R_1 remains 10).

In summary, in general, one should see the most improvement in using the optimal policy over simply giving priority to queue 1 if β_2 is large relative to β_1, and R_2 is close to R_1. To get an idea of the order of the maximum possible improvement (at least in this system), set $\beta_1 = 0$, $\beta_2 = 10$, and $R_2 = 9.99$. Here, $\rho_1 = 0.0832$, while ρ^* (the optimal policy gives priority to queue 2) is equal to 0.0945, an improvement of 13.6 percent. In the next section we will see that the improvements may be even more dramatic in the holding cost model.

4.2 The holding costs model

Here, we would like to again demonstrate the importance of taking abandonments into account, beyond through the index $h_i + \beta_i P_i$. We begin with a system that is almost symmetric. Let $\lambda_1 = \lambda_2 = 2$, $\mu = 4$, $h_1 = 1$, $h_2 = 0.99$, and $P_1 = P_2 = 1$. Note the loss of a customer in either queue is equally costly and the holding costs are close. Queue 1 will get priority according to our index, and Theorem 3.5 tells us that this policy is optimal if $\beta_2 \geq \beta_1$. If this condition is violated, then giving priority to queue 1 may yield poor performance. The intuition behind this is that if $\beta_1 > \beta_2$, then the higher rate of abandonments at queue 1 may mean that giving priority to queue 1 is simply too greedy.

To see this, we set $\beta_2 = 0$ and varied β_1, with the results in Table 4. (In Tables 4, 5, and 6, the final column gives the form of the optimal policy. The priority policies are

Table 4 Holding costs model, varying β_1

β_1	ρ_1	ρ^*	% increase	Policy
0.1	9.09	5.28	72.2	P2
0.2	8.38	3.94	112.7	P2
0.5	6.72	2.69	149.8	P2
1.0	5.00	2.08	140.4	P2
2.0	3.40	1.66	104.8	P2
4.0	2.32	1.38	68.1	P2
10.0	1.56	1.17	33.3	P2
100.0	1.05	1.01	4.0	P2

Table 5 Holding costs model, varying h_2

h_2	ρ_1	ρ^*	% increase	Policy
0.9	6.17	2.60	137.3	P2
0.8	5.56	2.50	122.4	P2
0.7	4.95	2.40	106.3	P2
0.6	4.34	2.30	88.7	P2
0.5	3.73	2.20	69.5	P2
0.4	3.12	2.08	50.0	DT
0.3	2.50	1.92	30.2	DT
0.2	1.89	1.69	11.8	DT
0.1	1.28	1.28	0	P1

Table 6 Holding costs model, varying β_2

β_2	ρ_1	ρ^*	% increase	Policy
0.1	3.69	2.49	48.2	P2
0.2	2.89	2.35	23.0	P2
0.3	2.49	2.24	11.2	P2
0.4	2.24	2.15	4.2	P2

is in the previous subsection, with the addition that DT denotes that the optimal policy is to give priority to class 2 if either the total number of customers in the system is above a threshold, or the number of class 1 customers is below a threshold.) Even with β_1 very small, there is a dramatic improvement by using the optimal policy (which gives priority to queue 2). The effect appears to be most prominent for moderate values of β_1 (relative to the service rate). At higher values of β_1, the improvement becomes less significant. The last row, $\beta_1 = 100.0$, suggests that a customer arriving to queue 1 either is serviced immediately or abandons, so there is little hope for the scheduling policy to have much impact.

As expected, this improvement is increasing with h_2 (Table 5 has results for varying h_2 with $\beta_1 = 0.5$ and $\beta_2 = 0$). Finally, to see that $\beta_2 = 0$ is not special, we fix $\beta_1 = 0.5$ and vary β_2 (Table 6), and we see that the improvement, which is still significant, decreases with increasing β_2 (as expected).

5 Conclusions/future work

In this paper we add abandonments to the classic (stochastic) scheduling model in a two-class service system. We do so under the two most common cost/reward structures; maximize rewards per service or minimize holding costs per customer per unit time. In each case the optimal scheduling rule, that holds without abandonments, no longer holds in general. Conditions for this simple priority rule to hold are provided. We also point to the fact that adding abandonments (in either case) causes several technical challenges. In particular, since the abandonment rate is not bounded, uniformization is not possible, and we must appeal to a continuous-time formulation of a Markov decision process instead of the discrete-time equivalent. Initially, this means the standard induction arguments cannot be applied. In the reward model, we use the continuous-time optimality equations and a sample path argument to show the result. However, even this method does not extend to the holding cost model. Only after a savvy use of truncation can the result be shown. As far as we know, this is the first time the continuous-time MDP formulation has been used to show structure in a queueing control problem.

Our numerical results highlight the point that a decision-maker that ignores the abandonments can significantly decrease the reward earned or increase the cost accrued. In the reward model, the added condition on the abandonment rates has an intuitive explanation and leads to a trade-off. The decision-maker needs to maximize rewards while minimizing the server idleness. When the rates are ordered in the same way as the rewards, both considerations can be handled simultaneously by prioritizing that class.

Characterizing the optimal policy in general (when our policies do not hold) is of clear interest. We have attempted to prove structural results in this case (in particular monotonicity), but to this point, such results have been elusive. Even if one could not characterize the policy over the entire parameter space, it would be of interest to provide a sharp condition under which the modified c–μ rule is optimal. Our conjecture is that this sharp condition would not be a simple expression.

There are several extensions that could be handled in future work. Perhaps the most obvious one is to consider more than 2 customer classes. The difficulty with multiple classes (even 3) is that the MDP formulation becomes more difficult to handle. For example, in the rewards model with 3 classes, we conjecture that it is optimal to prioritize station 1 when $R_1 = \max\{R_1, R_2, R_3\}$ and $\beta_1 = \max\{\beta_1, \beta_2, \beta_3\}$. To show this, we would need to show that $R_1 - R_2 + u_\alpha(i-1, j, k) - u_\alpha(i, j-1, k) \geq 0$ and $R_1 - R_3 + u_\alpha(i-1, j, k) - u_\alpha(i, j, k-1) \geq 0$. A sample path argument might do it but would be more tedious. In the holding cost case, we believe an analogous result holds, but since even the two-class case requires some adjustment to the truncation, the multiple class case seems unlikely to be a simple extension.

A second direction for examination is that of multiple servers. In the case of collaboration (when several servers can be assigned to the same customer), it seems that the current analysis holds. When servers cannot collaborate and there are but two classes, we believe the servers should avoid idling when the system state is close to the boundary, but again, the current insights hint toward what is optimal. The case of multiple servers and multiple customer classes is beyond the scope of this study and is still open.

Finally, we would like to point out that there are several other minor extensions. We have assumed that the service rates of each class are the same; that is, jobs assigned to the server in question are somewhat similar. We have also assumed that customers that are in service can abandon. This is akin to order cancellations or hang-ups after service has begun. It is our belief that in each case, each of these extensions make the problems far more tedious but do not add significantly to the insights provided here. We leave them for future research.

Acknowledgements The work of the first author is supported by the Natural Sciences and Research Council of Canada. The work of the third author is supported by the National Science Foundation under Grant Nos. CMMI-0900460 and CMMI-0826255. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Appendix

In this section we show that the optimality equations have a solution. Define $q(y|x,a)$ to be the rate at which a process leaves state x and goes to y given that action a is chosen. Recall for a continuous time Markov chain, $-q(x|x,a)$ is the rate at which a Markov process leaves state x given that action a is chosen. Denote the reward/cost rate in state x when using action a by $c(x,a)$. Let $q(x) := \sup\{-q(x|x,a) : a \in A(i)\}$. The following set of assumptions appear as Assumptions A, B, and C in [12]. Note that we are not making these assumptions in our work, rather we show that they all hold under our previously stated assumptions on the system.

Assumption A There exists a sequence $\{\mathbb{X}_m, m \geq 1\}$ of subsets of $\mathbb{X}$, a nondecreasing function $w \geq 1$ on $\mathbb{X}$, and constants, $b \geq 0$ and $c \neq 0$ such that

1. $\mathbb{X}_m \uparrow \mathbb{X}$ and $\sup\{q(x) : x \in \mathbb{X}_m\} < \infty$ for each $m \geq 1$;
2. $\inf\{w(x) : x \notin \mathbb{X}_m\} \to \infty$ as $m \to \infty$; and
3. $\sum_{y \in \mathbb{X}} w(y) q(y|x,a) \leq c w(x) + b$.

Assumption B

1. For every $(x,a) \in \{(y,a) : y \in \mathbb{X} \text{ and } a \in A(x)\}$ and some constant $M > 0$, $|c(x,a)| \leq M w(x)$, where $A(x)$ is the set of available actions in state x and w comes from Assumption A.
2. The discount factor $\alpha > 0$ is such that $\alpha > c$, where c is defined in Assumption A(3).

Assumption C

1. The action set $A(x)$ is compact for each $x \in \mathbb{X}$.
2. The functions $c(x,a)$, $q(y|x,a)$, and $\sum_{y \in \mathbb{X}} w(y) q(y|x,a)$ are all continuous in $a \in A(x)$ for each fixed $x, y \in \mathbb{X}$.
3. There exists a nonnegative function w' on $\mathbb{X}$ and constants $c' > 0$, $b' \geq 0$, and $M' > 0$ such that
   (a) $q(x) w(x) \leq M' w'(x)$, and

(b) for all (x, a),

$$\sum_{y \in \mathbb{X}} w'(y) q(y|x, a) \leq c' w'(x) + b'.$$

Lemma 6.1 *Suppose* $\alpha > \min\{\beta_1, \beta_2\}$ *and let*

$$D := \begin{cases} \mu \max\{R_1, R_2\} & \textit{for the reward model,} \\ \max\{h_1 + \beta_1 P_1, h_2 + \beta_2 P_2\} & \textit{for the holding cost model.} \end{cases}$$

In either the reward or holding cost models, Assumptions A, B, and C are satisfied with $\mathbb{X}_m = \{(i, j) | 0 \leq i, j \leq m\}$, $b = (\lambda_1 + \lambda_2) D + (\min\{\beta_1, \beta_2\})(\max\{D, 1\})$, $c = \min\{\beta_1, \beta_2\}$, *and* $w(i, j) := (i + j) D + \max\{D, 1\}$.

Proof We prove the result in the holding cost model. The reward model is analogous. To ease notation, let $\underline{\beta} := \min\{\beta_1, \beta_2\}$. Trivially, $\mathbb{X}_m \uparrow \mathbb{Z}^+ \times \mathbb{Z}^+$ as $m \uparrow \infty$; Assumption A(1) holds. Note that $w(i, j) \geq c((i, j), a)$. Of course, the fact that $w(i, j)$ is lower-bounded by $2mD + 1$ for $(i, j) \notin \mathbb{X}_m$ implies that Assumption A(2) holds. Note that, for $a = 1, 2$ (where the server will serve),

$$\begin{aligned} &\lambda_1 w(i + 1, j) + \lambda_2 w(i, j + 1) + \mu w\big(i - (2 - a), j + (1 - a)\big) + i\beta_1 w(i - 1, j) \\ &\quad + j\beta_2 w(i, j - 1) - (\lambda_1 + \lambda_2 + \mu + i\beta_1 + j\beta_2) w(i, j) \\ &= [\lambda_1 + \lambda_2 - \mu - i\beta_1 - j\beta_2] D \leq -\underline{\beta} w(i, j) + b, \end{aligned}$$

and Assumption A(3) is satisfied, as desired. Assumption B(1) is satisfied trivially, and Assumption B(2) holds by assumption. Since the action set is finite, the compactness and continuity conditions of Assumptions C(1) and C(2) are also trivial. It remains to consider Assumption C(3). Let

$$\begin{aligned} q(i, j) &:= \lambda_1 + \lambda_2 + \mu + i\beta_1 + j\beta_2 \\ &\leq \lambda_1 + \lambda_2 + \mu + (i + j) \max\{\beta_1, \beta_2\}. \end{aligned}$$

Define

$$\begin{aligned} w'(i, j) &:= \big((i + j) D + \max\{D, 1\}\big)\big[\lambda_1 + \lambda_2 + \mu + (i + j) \max\{\beta_1, \beta_2\}\big] \\ &= \big((i + j) D + \max\{D, 1\}\big)\big[\lambda_1 + \lambda_2 + \mu + (i + j)\overline{\beta}\big], \end{aligned}$$

where $\overline{\beta}$ is the maximal abandonment rate. We have $q(i, j) w(i, j) \leq M' w'(i, j)$ $(M' = 1)$. Moreover,

$$\begin{aligned} B(i, j) &:= \lambda_1 w'(i + 1, j) + \lambda_2 w'(i, j + 1) + \mu w'\big(i - (2 - a), j + 1 - a\big) \\ &\quad + i\beta_1 w'(i - 1, j) + j\beta_2 w'(i, j - 1) - (\lambda_1 + \lambda_2 + \mu + i\beta_1 + j\beta_2) w'(i, j) \\ &= (\lambda_1 + \lambda_2)\big(\big[\lambda_1 + \lambda_2 + \mu + 2(i + j)\overline{\beta}\big] D + \overline{\beta} D + \overline{\beta}\big(\max\{D, 1\}\big)\big) \\ &\quad - [\mu + i\beta_1 + j\beta_2]\big(\big[\lambda_1 + \lambda_2 + \mu + 2(i + j)\overline{\beta}\big] D + \overline{\beta} D - \overline{\beta}\big(\max\{D, 1\}\big)\big). \end{aligned}$$

Without loss of generality, assume that $D \ge 1$ (otherwise the constant b' becomes slightly more complicated). A little algebra yields

$$\begin{aligned}
B(i,j) &\le (\lambda_1+\lambda_2)\big(\big[\lambda_1+\lambda_2+\mu+2(i+j)\overline{\beta}\big]D+2\overline{\beta}D\big)\\
&= (\lambda_1+\lambda_2)\big(w'(i,j)-\big[\lambda_1+\lambda_2+\mu+(i+j)\overline{\beta}\big](i+j)D+2\overline{\beta}D\big)\\
&\quad + (\lambda_1+\lambda_2)\big[(i+j)\overline{\beta}\big]D\\
&\le (\lambda_1+\lambda_2)\big(w'(i,j)+2\overline{\beta}D\big).
\end{aligned}$$

Thus, Assumption C(3) holds with $c'=\lambda_1+\lambda_2$ and $b'=(\lambda_1+\lambda_2)2\overline{\beta}D$, and the proof is complete. □

Proof of Theorem 2.4 Given Lemma 6.1, the result is an immediate consequence of Theorem 3.2 of [12]. □

To prove Theorem 2.5, we proceed in much the same as in the discounted cost case. The following appears as Assumption A* in [12].

Assumption A* Assumptions A(1) and A(2) hold, and there exists a finite set $G \subset \mathbb{X}$, $b \ge 0$, and $c > 0$ such that

$$\sum_{y\in\mathbb{X}} w(y)q(y|x,a) \le -cw(x) + 1_{\{x\in G\}}b. \tag{6.1}$$

Lemma 6.2 *Assumption A* holds for w and w^2.*

Proof Let I' be the smallest integer such that for all $(i+j) \ge I'$, we have $(\underline{\beta}/2)w(i,j) \ge (\lambda_1+\lambda_2-\mu)D+\underline{\beta}\max\{D,1\}$ and define $\varphi = w(i,j)$ when $i+j=I'$. Recall from the proof of Lemma 6.1 that the left-hand side of (6.1) is bounded by

$$\begin{aligned}
&[\lambda_1+\lambda_2-\mu-i\beta_1-j\beta_2]D\\
&\quad\le (\lambda_1+\lambda_2-\mu)D-\underline{\beta}(i+j)D\\
&\quad= (\lambda_1+\lambda_2-\mu)D+\underline{\beta}\max\{D,1\}-\underline{\beta}\big[(i+j)D+\max\{D,1\}\big]\\
&\quad= (\lambda_1+\lambda_2-\mu)D+\underline{\beta}\max\{D,1\}-\underline{\beta}w(i,j)\\
&\quad= (\lambda_1+\lambda_2-\mu)D+\underline{\beta}\max\{D,1\}-(\underline{\beta}/2)w(i,j)-(\underline{\beta}/2)w(i,j)\\
&\quad\le -(\underline{\beta}/2)w(i,j)+\varphi 1_{\{(i+j)\le I'\}},
\end{aligned}$$

where the last inequality holds by assumption and completes the proof.

Consider now w^2. The left-hand side of (6.1), with the addition of $cw(i,j)$ for some $c>0$ (to be defined later), can be written (for $(i+j)\ge 1$)

$$\begin{aligned}
&(\lambda_1+\lambda_2)\big[(i+j+1)D+1\big]^2+(\mu+i\beta_1+j\beta_2)\big[(i+j-1)D+1\big]^2\\
&\quad+\big[c-(\lambda_1+\lambda_2+\mu+i\beta_1+j\beta_2)\big]\big[(i+j)D+1\big]^2
\end{aligned}$$

$$
\begin{aligned}
&= (\lambda_1+\lambda_2)\big[\big((i+j)D+1\big)+D\big]^2 + (\mu+i\beta_1+j\beta_2)\big[\big((i+j)D+1\big)-D\big]^2 \\
&\quad + \big[c-(\lambda_1+\lambda_2+\mu+i\beta_1+j\beta_2)\big]\big[(i+j)D+1\big]^2 \\
&= (\lambda_1+\lambda_2)\big[2D\big((i+j)D+1\big)+D^2\big] \\
&\quad + (\mu+i\beta_1+j\beta_2)\big[-2D\big((i+j)D+1\big)+D^2\big]+c\big[(i+j)D+1\big]^2 \\
&\le (\lambda_1+\lambda_2)\big[2D\big((i+j)D+1\big)+D^2\big] \\
&\quad + \big(\mu+(i+j)\underline{\beta}\big)\big[-2D\big((i+j)D+1\big)+D^2\big]+c\big[(i+j)D+1\big]^2. \qquad (6.2)
\end{aligned}
$$

Consider the quadratic term $(i+j)^2D^2(c-2\underline{\beta})$. Thus, for $c<2\underline{\beta}$ and $(i+j)$ sufficiently large, the expression in (6.2) is nonpositive. The quadratic term dominates. Let I' be such that the expression in (6.2) is nonpositive for $(i+j)\ge I'$ and denote the maximum of this expression for $(i+j)\le I'$ by φ. The result follows. □

Proof of Theorem 2.5 The fact that w^2 satisfies (6.1) along with the irreducibility implies that Assumption D of [12] holds (see the comments following Proposition 4.2 of [12]). The theorem is now a direct application of Theorem 4.1 of [12]. □

References

1. Aksin, Z., Armony, M., Mehrotra, V.: The modern call-center: a multi-disciplinary perspective on operations management research. Prod. Oper. Manag. **16**, 665–668 (2007)
2. Argon, N.T., Ziya, S., Righter, R.: Scheduling impatient jobs in a clearing system with insights on patient triage in mass casualty incidents. Probab. Eng. Inf. Sci. **22**(3), 301–332 (2008)
3. Armony, M., Maglaras, C.: Contact centers with a call-back option and real-time delay information. Oper. Res. **52**(4), 527–545 (2004)
4. Armony, M., Maglaras, C.: On customer contact centers with a call-back option: customer decisions, routing rules and system design. Oper. Res. **52**(2), 271–292 (2004)
5. Armony, M., Shimkin, N., Whitt, W.: The impact of delay announcements in many-server queues with abandonment. Oper. Res. **57**, 66–81 (2009)
6. Atar, R., Giat, C., Shimkin, N.: The $c\mu/\theta$ rule for many-server queues with abandonment. Preprint (2010)
7. Buyukkoc, C., Varaiya, P., Walrand, J.: The $c\mu$-rule revisited. Adv. Appl. Probab. **17**(1), 237–238 (1985)
8. Gans, N., Koole, G., Mandelbaum, A.: Telephone call centers: tutorial, review and research prospects. Manuf. Serv. Oper. Manag. **5**(2), 79–141 (2003)
9. Gayon, J.-P., Benjaafar, S., de Véricourt, F.: Using imperfect advance demand information in production-inventory systems with multiple customer classes. Manuf. Serv. Oper. Manag. **11**(1), 128–143 (2009)
10. Ghamami, S., Ward, A.: Dynamic scheduling of an N-system with reneging. Preprint (2010)
11. Guo, X., Hernández-Lerma, O.: Continuous-time controlled Markov chains. Ann. Appl. Probab. **13**(1), 363–388 (2003)
12. Guo, X., Hernández-Lerma, O., Prieto-Rumeau, T.: A survey of recent results on continuous-time Markov decision processes. Top **14**(2), 177–257 (2006)
13. Harrison, J., Zeevi, A.: Dynamic scheduling of a multiclass queue in the Halfin and Whitt heavy traffic regime. Oper. Res. **52**, 243–257 (2004)
14. Iravani, F., Balcıoğlu, B.: On priority queues with impatient customers. Queueing Syst. **58**, 239–260 (2008)
15. Koçağa, Y., Ward, A.: Admission control for a multi-server queue with abandonment. Preprint (2010)

16. Lippman, S.: Applying a new device in the optimization of exponential queueing systems. Oper. Res. **23**(4), 687–710 (1975)
17. Nain, P.: Interchange arguments for classical scheduling problems in queues. Syst. Control Lett. **12**, 177–184 (1989)
18. Prieto-Rumeau, T., Hernández-Lerma, O.: A unified approach to continuous-time discounted Markov control processes. Morfismos **10**(1), 1–40 (2006)
19. Puterman, M.: Markov Decision Processes: Discrete Stochastic Dynamic Programming. Wiley Series in Probability and Mathematical Statistics. Wiley, New York (1994)
20. Serfozo, R.: An equivalence between continuous and discrete time Markov decision processes. Oper. Res. **27**(3), 616–620 (1978)
21. Tezcan, T., Dai, J.: Dynamic control of N-systems with many servers. asymptotic optimality of a static priority policy in heavy traffic. Preprint (2010)
22. Ward, A.R., Glynn, P.W.: A diffusion approximation for a Markovian queue with reneging. Queueing Syst. **43**(1), 103–128 (2003)
23. Ward, A.R., Glynn, P.W.: A diffusion approximation for a GI/GI/1 queue with balking or reneging. Queueing Syst. **50**(4), 371–400 (2005)